

Construction of Weights for the 1996  
Medical Expenditure Panel Survey  
Insurance Component List Sample

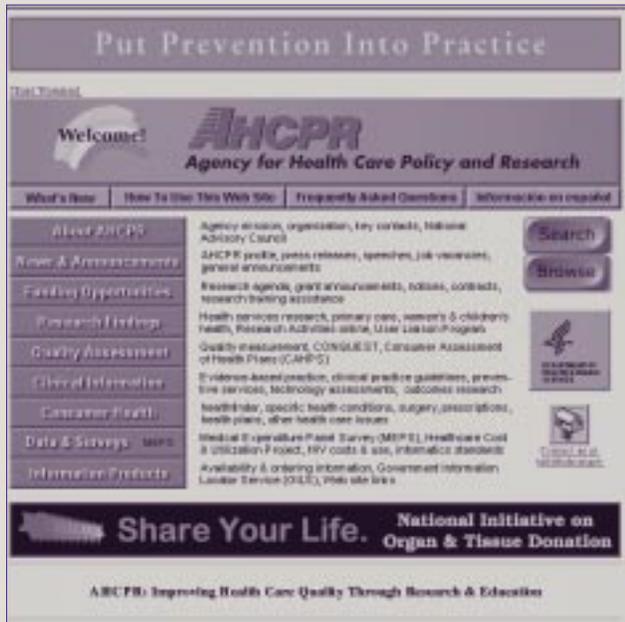
# Methodology

## Report 8



U.S. Department of Health and Human Services  
Public Health Service  
Agency for Health Care Policy and Research

# Health Care Information and Electronic Ordering Through the AHCPR Web Site



The Agency for Health Care Policy and Research's Web site—<http://www.ahcpr.gov/>—makes practical, science-based health care information available in one convenient place.

Buttons correspond to major categories of Web site information, including funding opportunities, research findings, quality assessments, clinical information, consumer health, and data.

The Web site features an Electronic Catalog to the more than 450 information products generated by AHCPR, with information on how to obtain these resources. Many information products have an electronic ordering form and are mailed free of charge from the AHCPR Clearinghouse within 5 working days.

<http://www.ahcpr.gov/>

## Abstract

The Medical Expenditure Panel Survey (MEPS) is the third in a series of nationally representative surveys of medical care use and expenditures sponsored by the Agency for Health Care Policy and Research (AHCPR). MEPS comprises four component surveys. The Insurance Component (IC) collects employment-related health insurance information, such as premiums and types of plans offered. Respondent characteristics—such as size of business, employee characteristics, and type of industry—also are collected. The IC sample comprises a household sample component (linked to the MEPS household survey) and a list sample. The list sample is an independently selected random sample of governments and private-sector establishments. Its

purpose is to make national and State estimates of employer insurance characteristics, costs, and the numbers of employees enrolled. This report describes the overall response rates for the list sample and the process used to correct the weights for respondents. The weights are corrected in order to adequately represent all nonrespondents and all important subgroups, especially subgroups that may have had different response rates from the average for the survey.

## Suggested citation

Sommers JP. Construction of weights for the 1996 Medical Expenditure Panel Survey Insurance Component list sample. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 8. AHCPR Pub. No. 00-0005.

Construction of Weights for the 1996  
Medical Expenditure Panel Survey  
Insurance Component List Sample

# Methodology

## Report 8



**U.S. Department of Health and Human Services**  
Public Health Service  
Agency for Health Care Policy and Research  
AHCPR Pub. No. 00-0005  
November 1999

# The Medical Expenditure Panel Survey (MEPS)

## Background

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS also includes a nationally representative survey of nursing homes and their residents. MEPS is cosponsored by the Agency for Health Care Policy and Research (AHCPR) and the National Center for Health Statistics (NCHS).

MEPS comprises four component surveys: the Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHCPR on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES) was conducted in 1977, the National Medical Expenditure Survey (NMES) in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features

include linkage with the National Health Interview Survey (NHIS), from which the sample for the MEPS HC is drawn, and enhanced longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

## Household Component

The MEPS HC, a nationally representative survey of the U.S. civilian noninstitutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for 2 calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with oversampling of Hispanics and blacks.

## Medical Provider Component

The MEPS MPC supplements and validates information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the

HC. Also included in the MPC are all office-based physicians:

- Providing care for HC respondents receiving Medicaid.
- Associated with a 75-percent sample of households receiving care through an HMO (health maintenance organization) or managed care plan.
- Associated with a 25-percent sample of the remaining households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9 (9th Revision, International Classification of Diseases) and DSM-IV (Fourth Edition, *Diagnostic and Statistical Manual of Mental Disorders*).
- Physician procedure codes classified by CPT-4 (Current Procedural Terminology, Version 4).
- Inpatient stay codes classified by DRG (diagnosis-related group).
- Prescriptions coded by national drug code (NDC), medication names, strength, and quantity dispensed.
- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials.

## Insurance Component

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, and employer characteristics.

Establishments participating in the MEPS IC are selected through four sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.
- A Bureau of the Census list frame of private-sector business establishments.
- The Census of Governments from the Bureau of the Census.

- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and other insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

The MEPS IC is an annual panel survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone followup for nonrespondents.

## Nursing Home Component

The 1996 MEPS NHC was a survey of nursing homes and persons residing in or admitted to nursing homes at any time during calendar year 1996. The NHC gathered information on the demographic characteristics, residence history, health and functional status, use of services, use of prescription medications, and health care expenditures of nursing home residents. Nursing home administrators and designated staff also provided information on facility size, ownership, certification status, services provided, revenues and expenses, and other facility characteristics. Data on the income, assets, family relationships, and caregiving services for sampled nursing home residents were obtained from next-of-kin or other knowledgeable persons in the community.

The 1996 MEPS NHC sample was selected using a two-stage stratified probability design. In the first stage, facilities were selected; in the second stage, facility residents were sampled, selecting both persons in residence on January 1, 1996, and those admitted during the period January 1 through December 31.

The sampling frame for facilities was derived from the National Health Provider Inventory, which is updated periodically by NCHS. The MEPS NHC data were collected in person in three rounds of data collection over a 1½-year period using the CAPI system. Community data were collected by telephone using computer-assisted telephone interviewing (CATI) technology. At the end of three rounds of data collection,

the sample consisted of 815 responding facilities, 3,209 residents in the facility on January 1, and 2,690 eligible residents admitted during 1996.

## Survey Management

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files.

Printed documents and CD-ROMs are available through the AHCPR Publications Clearinghouse. Write or call:

AHCPR Publications Clearinghouse  
Attn: (publication number)  
P.O. Box 8547  
Silver Spring, MD 20907  
800-358-9295  
410-381-3150 (callers outside the United States only)  
888-586-6340 (toll-free TDD service; hearing impaired only)

Be sure to specify the AHCPR number of the document or CD-ROM you are requesting. Selected electronic files are available through the Internet on the AHCPR Web site:

<http://www.ahcpr.gov/>

On the AHCPR Web site, under Data and Surveys, click the MEPS icon.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Health Care Policy and Research, 2101 East Jefferson Street, Suite 500, Rockville, MD 20852 (301-594-3075).

## Table of Contents

Introduction . . . . .	1
Response Rates. . . . .	1
Private-Sector Weights . . . . .	3
Government Weights. . . . .	7
Representative Results. . . . .	7
References . . . . .	8
Appendix. Definition of Selected Variables. . . . .	9

### Tables showing information on the 1996 MEPS Insurance Component list sample:

1. Private-sector response . . . . .	2
2. Government response . . . . .	3
3. Selected statistics. . . . .	7

# Construction of Weights for the 1996 Medical Expenditure Panel Survey Insurance Component List Sample

by John Paul Sommers, Ph.D., Agency for Health Care Policy and Research

## Introduction

The Medical Expenditure Panel Survey (MEPS) Insurance Component (IC) is a survey of employers, unions, and other providers of health insurance. It is sponsored by the Agency for Health Care Policy and Research (AHCPR) and is conducted by the U.S. Bureau of the Census. The IC sample is composed of two sets of employers and other providers of health insurance.

The first part of the IC sample is the list sample, an independently selected random sample of governments and private-sector establishments. The purpose of this sample of just over 40,000 employers is to make national and State estimates of employer insurance characteristics, their costs, and the numbers of employees enrolled (Sommers, 1999).

The second part of the IC sample, the household sample component, serves a different purpose: it is person oriented, not employer oriented. The two samples are collected together because of the similarity in data collected from similar types of respondents.

The IC household sample component is the set of employers and other providers of health insurance for persons who were members of the sample of the MEPS Household Component (HC), a household survey. Respondents in this part of the IC sample serve as proxy respondents, giving employer and insurance information for household members who are part of the HC sample. The data for household members collected during the IC are attached as data elements to other information collected directly from the household respondents. Weights derived for the data from the household sample component of the IC are thus dependent on the sample design of the HC (J. Cohen, 1997; S. Cohen, 1997). The development of weights for the MEPS HC will be presented in a separate document.

This report briefly describes the overall response rates for the list sample and the process used to correct the weights for respondents. The weights are corrected in order to adequately represent all nonrespondents and

all important subgroups, especially subgroups that may have had different response rates from the average for the survey.

## Response Rates

In this section, various response rates are given for the list sample. Response rates for subgroups of the population that may have characteristics different from the average are especially important. Their importance is discussed in later sections describing the construction of weights for sampling units. Before the rates are presented, it is necessary to discuss the data collection methodology used for the IC. This affects the definition of a response and the method selected to create weights.

The IC sample is composed of governments and private-sector establishments. Governments were selected from the Census of Governments, and private-sector establishments from the Census Bureau's Standard Statistical Establishment List (Sommers, 1999). For all samples except State and very large local governments, each sample unit was prescreened. The purpose of this step was to determine a point of contact for data collection and determine whether the employer offered insurance to its employees. If the employer did not offer insurance, a small number of questions were administered and, with this done, the employer was considered a respondent. This was a quick and inexpensive method to collect the necessary data from the large number of employers that do not offer health insurance to their employees.

Employers that offered insurance were asked several brief questions and then mailed a questionnaire about their insurance. If they failed to return the mail questionnaire, an attempt was made to collect the information by telephone. For the purpose of this survey, employers that offered insurance were considered respondents only if they answered key questions on their health insurance. Those that did not were considered nonrespondents.

If no contact was made by telephone, a questionnaire was mailed. If the questionnaire was not returned, another attempt was made to collect information by telephone. Any employer from this group that responded by mail or telephone was a respondent. Note that for this group whether a nonrespondent had insurance was unknown.

Table 1 shows the private-sector sample by inscope and respondent status. Out-of-scope private-sector establishments were those that had gone out of business or consisted of a self-employed person with no other employees.

Although an establishment with insurance that does not answer questions about the plans offered is considered a nonresponse for the IC, one can argue that

it provided certain valuable information. If one assumed that all cases were successfully prescreened, then the entire sample size could be used to estimate the percentage of establishments that offer insurance. As will be shown in the section describing the weighting methodology used, this partial information can be used as a part of the correction for nonresponse.

Table 2 shows the government sample by inscope and respondent status. Private-sector and government employers are shown separately because of the differences in response rates and other features. To be out of scope, a government must have ceased to function. To be a respondent, a government had to answer questions similar to those for a private-sector respondent.

**Table 1. Private-sector response in the 1996 Medical Expenditure Panel Survey Insurance Component list sample**

Status	Number
Total sample	37,710
Out of scope	2,567
In scope	35,143
Respondents (66.1 percent of inscope sample)	23,232
Completed prescreener	20,003
Offered insurance	13,622
Did not offer insurance	6,381
Did not complete prescreener	3,229
Nonrespondents (33.9 percent of inscope sample)	11,911
Completed prescreener, offered insurance	5,061
Incomplete prescreener, unknown insurance status	6,850

**Source:** Center for Cost and Financing Studies, Agency for Health Care Policy and Research: Medical Expenditure Panel Survey Insurance Component, 1996.

**Table 2. Government response in the 1996 Medical Expenditure Panel Survey Insurance Component list sample**

Status	Number
Total sample	2,649
Out of scope	20
In scope	2,629
Respondents (84.6 percent of inscope sample)	2,224
Completed prescreener	1,049
Offered insurance	1,049
Did not offer insurance	0
Did not complete prescreener	1,175
Nonrespondents (15.4 percent of inscope sample)	405
Completed prescreener, offered insurance	151
Incomplete prescreener, unknown insurance status	254

**Source:** Center for Cost and Financing Studies, Agency for Health Care Policy and Research: Medical Expenditure Panel Survey Insurance Component, 1996.

## Private-Sector Weights

### Motivation for Weighting

Because the sample units were selected with unequal probabilities and because of nonresponse in the data collection, sampling weights must be calculated for each unit in the original sample and the weights of respondents must be adjusted for nonresponse (Kish, 1965). To motivate this adjustment, it is assumed that  $p_i$  is the probability of selection of the  $i$ th unit in the sample, and  $wt_i = 1/p_i$ , the initial weight. Then

$$E\left[\sum_{i \in S} wt_i x_i\right] = \sum_{i \in U} p_i wt_i x_i = \sum_{i \in U} \left[\frac{p_i}{p_i}\right] x_i = \sum_{i \in U} x_i$$

where  $S$  and  $U$  are the set of sampled units and set of units in the universe, respectively.

A method of making an unbiased estimate of a total from the sample is to use the weighted sum from the sample as the estimate. If there is nonresponse—if, for example, some of the  $x$  values are missing—one would try to adjust sampling weights for the values of respondents in cells for which the responding  $x$ 's have the same or very similar expected values to those of the nonresponding  $x$ 's. Thus, the final objective is to have the sum of the new adjusted weights of respondents over cells with various characteristics equal to the totals from the original sample or some other set of known control totals for the population.

## Classification of Establishments

Cells defined by size of establishment, size of firm, and State were used to create cells for weighting. These same characteristics had been used for sample selection (Sommers, 1999). It was first thought that the original cell classification of establishments would be used. However, the following factors led to a decision to update the classification:

- The frame used for sampling was preliminary for 1996. The Census Bureau had since produced a final listing of establishments for March 1996.
- The reported and verified employment sizes of establishments sometimes differed considerably from the employment sizes on the frame.

A scheme was developed to update the size classification. Otherwise, a unit with a frame employment of 2 but a reported establishment employment of 2,000 would cause extreme results during the poststratification process (described later in this report).

Establishments were divided into the six size classes used for sample selection. (See Sommers, 1999, and the appendix.) If the reported employment would have placed an establishment more than one cell away from its original frame-defined cell, then the establishment was reclassified. These establishments were considered to have been mistakes on the frame. Establishments that moved only one cell were allowed because the date of the frame employment was early March and the reference date in the survey was July of the same year. Since there is seasonality and growth in employment, with March being a relatively low month for employment and July high, it seemed that this change should be allowed (Bureau of Labor Statistics, 1996). Approximately 5.6 percent of the private-sector sample changed its establishment size cell either up or down at least two classes to correspond to the reported employment. The remaining establishments kept their original frame classification.

## Prescreener Respondent Weight Adjustments

As discussed earlier, during data collection a telephone prescreener was conducted. Employers that reported during this prescreener that they did not offer insurance to their employees were considered respondents. Those that did offer health insurance were not considered respondents unless they provided further information on the insurance either by mail or telephone. As shown in Table 1, 5,061 private-sector employers (14.4 percent of the sample) were prescreened, had insurance, and later were classified as nonrespondents. However, after the prescreener, 25,064, or 71.3 percent of the inscope sample, had answered whether they provided insurance. It is of utmost importance that this information be retained in the weights after adjustment for nonresponse. Thus, the first step in weighting was to adjust the weights of respondents that provided insurance and were successfully prescreened so that their adjusted total weight would be equal to the total weight for all employers who were successfully prescreened and provided insurance to their employees.

First, weights for the respondents that offered insurance and had also responded to the prescreener were adjusted so that their total weight equaled the weight of all respondents to the prescreener that offered insurance for a large number of cells. Among the variables chosen to define these important cells were:

- A variable with 41 levels that was defined by crossing groups of States with a variable with 8 levels based on employment size of the establishment and that of the enterprise as listed on the frame (Sommers, 1999).
- Whether the establishment was from a firm with multiple establishments.
- The industry of the establishment.

Because crossing all these variables divides the sample into too many cells, a method called raking was employed to create an adjustment that retained the sum of the weights for as many cells as possible.

In an ideal world, if the value of the  $i$ th unit's initial weight were  $w_{t_i}$ , the adjusted weight for responding

units within each response cell C, where CR is the set of respondents in C, would be

$$adjwt_i = \frac{wt_i \sum_{j \in C} wt_j}{\sum_{j \in CR} wt_j}$$

The adjusted weights for nonrespondents would be set to zero. This would be done for all cells C and CR defined by the cross of all the important variables.

Because the number of cells defined by the cross of all levels of these variables was very large, the sum of the respondent weights for each cell created by these cross-classifications was not made equal to the population total. Instead, the sums of the respondent weights for all cells defined by any one of the variables alone were made to sum to the totals for respondents and nonrespondents in the cell. Thus, the sum of the respondents' weights equaled the total sample weights for all the marginal cells defined by these variables.

The method can be illustrated with an example. Suppose one were concerned only with two variables: (1) type of firm, with 2 levels—single establishment and multi establishment, and (2) industry, with 7 levels. The sum of weights of respondents in the 14 cells defined by a cross of the two variables (for instance, single *and* mining) is not made to equal the total of sampling weights for all establishments in each of the cells. Instead, the sum of the weights for respondents for each of the nine marginal cells defined by one of the individual variables (for instance, single *or* mining) is made to equal the total weights of all establishments in the cell. To accomplish this, one uses an iterative technique called raking, which is explained below (Madow, Olkin, and Rubin, 1983).

Assume  $wt_{ijk}$  is the sampling weight for the  $i$ th sample unit in the  $jk$ th cell. For example, the first variable could break the population by type of firm and the second by industry. One would adjust the weights to create a set of weights  $wt'_{ijk}$  such that

$$N_j = \sum_{i,k \in CR} wt'_{ijk} = \sum_{i,k \in C} wt_{ijk} \text{ and } N_k = \sum_{i,j \in CR} wt'_{ijk} = \sum_{i,j \in C} wt_{ijk}$$

To do this, one first creates values  $a_j$  such that

$$a_j = \frac{N_j}{\sum_{i,k \in CR} wt_{ijk}}$$

Then  $wt'_{ijk} = a_j * wt_{ijk}$  for all cases where the  $ijk$ th case is a respondent. Otherwise, the adjusted weight is 0.

One continues this process through all values of  $j$ . This makes the respondents' weights sum to the proper values for the first defining variable, but not the second. One then operates on the values of  $wt'_{ijk}$  to adjust them so that the cells defined by the second variable are equal. Of course, now this third set of weights may not sum properly for cells defined by the first variable. The cycle is repeated, readjusting each new set of weights by continuing to iterate through the process. Usually, within several cycles, the weights will converge so that for each cell defined by either of the two variables, the sum of the latest set of weights is a value very close to the required marginal total. This completes the raking process. The process can be applied to make weights sum to marginal cell totals for as many variables as one likes. In this example, two variables were used to create marginal totals and the weights were forced to add to the sample total for the type of firm and for each industry, but not necessarily by type of firm and industry.

In adjustments for the sampling weights for respondents offering insurance that also responded to the prescreener, three variables defined the marginal cells and their totals for this process:

- Eight combinations of establishment size and firm size within the 41 State sets. (Each of the 40 most populous States was a set, and the District of Columbia and the 10 least populous States were grouped together as a single set.)
- Multi and single establishment firms.
- Industry type.

Exact category definitions can be found in the appendix.

## Prescreener Nonrespondent Weight Adjustments

Weights for the remaining respondents from which no prescreener information was collected must be adjusted to account for nonrespondents that did not provide prescreener information. Because of the lack of the prescreener information, there is no partial response in this group. Establishments in this group could either provide or not provide insurance to their employees. This part of the sample adjustment is similar to that done for most surveys which have only one stage of data collection. A nonresponse correction adjusted the current set of weights of all 23,232 (6,381 + 13,622 + 3,229) complete respondents, regardless of their prescreener status, so that their weights after the correction also included the total sampling weight of the 6,850 nonrespondents that had an incomplete prescreener and unknown insurance status. The weights used for the respondents were those currently available. For the 6,381 prescreener respondents that did not offer insurance and the 3,229 respondents that did not answer the prescreener, these were the original sampling weights. For the prescreener respondents that offered insurance, they were the weights after the first adjustment, described above.

To make this second adjustment, raking was used. The variables and cells used were the same as those used for the previous correction for those nonrespondents that answered the prescreener and offered insurance.

## Poststratification of Private-Sector Weights

Correction of weights can be taken further using a technique called poststratification. This method is used when outside information is available about the population. The weighted sum of a variable can be thought of as an estimate of the population total for the variable. Ideally, the total weighted sum for each important subset of the population should equal the population for that subset. Controlling weights to population totals from another source is called

poststratification (Madow, Olkin, and Rubin, 1983; Skinner, Holt, and Smith, 1989).

The MEPS IC private-sector sample is selected from a list of establishments, the Standard Statistical Establishment List (SSEL) maintained by the Census Bureau (Kreps, Slater, and Plotkin, 1979; Sommers, 1999). This list contains each establishment that was in business during the year and its employment size as of March of the year. It is the source of official Census Bureau figures of the number and employment size of establishments in the United States. For this reason, it is very desirable that the estimates of employment obtained from the IC match the values from the SSEL for the total and a variety of subtotals.

Because establishments go out of business during the year, the list used for the IC represented a point in time, establishments in business as of a certain date, rather than the entire frame for the year. It was felt that this type of list, showing the entire population of workers at one instant, would give a better picture of the percentage of establishments in business at any point that were offering insurance and the numbers of employees in the work force offered insurance. It was felt that including all establishments in business during the year might skew the results. Small establishments tend not to offer health insurance and also go in and out of business more rapidly than larger ones. Thus to include all the small businesses from the entire year would give a higher estimate of the numbers of establishments not offering insurance to their employees and the number of employees not offered insurance. It also could tend to double-count employees who moved from closed establishments where they were likely not to have had insurance. Thus, one could get a higher proportional count of people not offered insurance.

The SSEL used to provide control totals for the IC was the list of establishments in business as of the end of the third quarter of 1996. Because of the retrospective time frame of the data collected, this information was available in late 1997, when weighting took place. This time frame also fit well with the timing of the data collected.

To implement this process, the same  $41 \times 8$  categories were used that were assigned during the

raking process. These cells were based on the establishment's location, firm size, and employment size. The size used was the one reported for the frame, not the size reported during the survey, except for establishments that were reassigned to new establishment size classes during the classification process described earlier. The size reported for the SSEL frame was used because the survey period was later in the year than the SSEL, and it was felt that the frame categories and employment sizes were the most comparable variables. The frame employment for the SSEL is for early March of 1996.

The weighted frame employment sizes from the responding sample establishments were obtained using the weights after the two other adjustments and the total employment on the frame for in-scope establishments were calculated for each of the 328 categories. The ratio of the frame employment size for March to the estimated employment size from the sample was calculated for each category. A new corrected weight for each respondent in the category was created by multiplying its category ratio times the current value of its corrected weight before poststratification. This procedure corrected the sample weights so that the sample employment for March equaled the frame employment. The collected employment estimate was somewhat higher than the frame employment because employment tends to have a lower seasonal component in March than in the later months of the year for which data were collected (Bureau of Labor Statistics, 1996).

## Government Weights

Data for governments were collected in the same manner as for the private sector: a prescreener phone contact was used as the first contact, and the process for adjusting weights parallels that for the private sector. The differences are in the definition and sizes of the cells used. Cells were determined by a cross of the 41 State sets (40 most populous States and the residual set of 10 States and the District of Columbia) and 5 classes based on size and type of government. The classes are listed in the appendix. These cells were used for two nonresponse adjustments and for poststratification of governments.

As with the private-sector weights, the last step of the process was a poststratification adjustment. Units within the  $41 \times 5$  cells were adjusted using the ratio of the total 1996 frame employment size for the cell divided by the weighted total 1996 employment size for the sample within the cell.

## Representative Results

Table 3 contains some typical results for the private sector along with their relative standard errors.

**Table 3. Selected statistics from the 1996 Medical Expenditure Panel Survey Insurance Component list sample**

Item	Estimate	Relative standard error (percent)
Total establishments	5,998,994	0.6
Total employment	103,846,469	1.5
Percent of establishments offering insurance	53.2	0.5
Average single premium per enrolled employee	\$1,996.90	0.8
Average single employee contribution	\$338.33	2.6

**Source:** Center for Cost and Financing Studies, Agency for Health Care Policy and Research: Medical Expenditure Panel Survey Insurance Component, 1996.

## References

Bureau of Labor Statistics (US). Current employment statistics. Washington. Series EEU00500005; 1996.

Cohen J. Design and methods of the Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS Methodology Report No. 1. AHCPR Pub. No. 97-0026.

Cohen S. Sample design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS Methodology Report No. 2. AHCPR Pub. No. 97-0027.

Kish L. Survey sampling. New York: John Wiley and Sons; 1965.

Kreps J, Slater CM, Plotkin MD. The Standard Statistical Establishment List Program. Washington: U.S. Bureau of the Census; 1979. Technical Paper No. 44.

Madow WG, Olkin I, Rubin DR. Incomplete data in sample surveys, volume 2: theory and bibliographies. New York: Academic Press; 1983.

Skinner CJ, Holt D, Smith TMF. Analysis of complex surveys. New York: John Wiley and Sons; 1989.

Sommers JP. List sample design of the 1996 Medical Expenditure Panel Survey Insurance Component. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 6. AHCPR Pub. No. 99-0037.

## Appendix. Definition of Selected Variables

### Industry type (Standard Industry Code)

- 1 Construction, SIC = 15-17
- 2 Mining, manufacturing, SIC = 10-14, 20-39
- 3 Transportation, communications, utilities, SIC = 40-49
- 4 Wholesale trade, SIC = 50-51
- 5 Agriculture, retail trade, SIC = 01-09, 52-59
- 6 Finance, insurance, real estate, SIC = 60-67
- 7 Services, SIC = 70-89
- 8 Unknown, SIC = missing

### Firm size

- S Less than 50 employees
- M 50-999 employees
- L 1,000 employees or more

### Establishment size

- 1 1-5 employees
- 2 6-24 employees
- 3 25-49 employees
- 4 50-249 employees
- 5 250-999 employees
- 6 1,000 employees or more

### Collapsed firm size × establishment size

- 1 Firm size = S, establishment size = 1
- 2 Firm size = S, establishment size = 2

- 3 Firm size = S, establishment size = 3
- 4 Firm size = M, establishment size = 1,2
- 5 Firm size = M, establishment size = 3,4
- 6 Firm size = L, establishment size = 1,2
- 7 Firm size = L, establishment size = 3,4
- 8 Firm size = M, L, establishment size = 5, 6

### Multi/single establishment firm

1 if multi, 0 otherwise

### State groupings

AK, DE, DC, ID, MT, NH, ND, RI, SD, VT, and WY are grouped together because of small sample sizes

All 40 other States stand on their own

### Government size groups

- 1 if a State government
- 2 if a local certainty
- 3 if a local noncertainty government and frame employment < 200
- 4 if a local noncertainty government and 199 < frame employment < 1,000
- 5 if a local noncertainty government and frame employment is 1,000 or more

**Note:** State governments were classified separately because there is only one per State (they are unique within each State).

**U.S. Department of Health and Human Services**  
Public Health Service  
Agency for Health Care Policy and Research  
2101 East Jefferson Street, Suite 501  
Rockville, MD 20852



AHCPR Pub. No. 00-0005  
November 1999